

Books

Arguing A.I.: The Battle for Twenty-First Century Science, Sam Williams, Random House, New York, 2002. 128 pp. (ISBN 0-8129-91180-X).

The study of intelligence, real and artificial, has always given rise to passionate philosophical and social debates. In general, these boil down to the two Big Questions: First, can we understand intelligence in purely mechanistic terms and use that understanding to build intelligent machines—perhaps even superhuman ones? And second, if we could do that, would it be a good thing to do? Would this be the next great step in evolution, or a stupid, perhaps suicidal act on the part of the human species? Or both?

These are two very different questions, although they are often muddled together in various ways. Obviously, the urgency of the second question depends in large part on your answer to the first one.

At present, we are witnessing a new flare-up in this long-running argument (or set of arguments). In 1999 Ray Kurzweil published his book, *The Age of Spiritual Machines*. In it, he argues that the exponential increase of computing power (Moore's Law) makes the arrival of superhuman intelligent machines inevitable and that this development is closer than we may think. Kurzweil views this as an exciting and positive development. Other authors have made similar arguments in the past, but none with Kurzweil's combination of solid credentials in artificial intelligence (AI) and flair for promoting his ideas.

Shortly after Kurzweil's book appeared, Bill Joy published the article "Why the Future Doesn't Need Us" in *Wired* magazine. Joy is a well-respected system developer, although he has never worked in AI. He more or less accepts Kurzweil's prediction that in-

telligent, self-replicating machines are coming soon, and he worries that these may easily escape our control, with dire results. (Joy is even more worried about genetic technology and self-replicating nanosystems, but that's another story.) The popular press has picked up on this colorful and entertaining disagreement, sometimes carelessly describing it as "the" AI debate.

In this slim book, author Sam Williams reports on this argument between Kurzweil and Joy and gives an account of the swirling intellectual currents that led to the current arguments.

Williams adds some additional viewpoints to round out the picture. Self-described "humanist" Jaron Lanier accuses AI of removing the human element from computing. In his view, this somehow makes AI responsible for bad user-interface design and bug-infested software. (This seems a bizarre suggestion, since some of the most important innovations in user-interface design and software-development tools have taken place in and around AI research groups.)

AI pioneer John McCarthy is also profiled. He states the viewpoint of many who actually work in the field of AI: faster machines may be necessary for human-level intelligence, but they are not sufficient. As a field, we still seem to be missing some key ideas, and nobody can predict how long it will be until all the missing pieces come together. Are we missing just one or two big ideas, or dozens of smaller ones? Nobody knows. But AI researchers with a realistic appreciation for the difficulties still facing the field do not spend a lot of time worrying about the sudden

©Copyright 2002 by International Business Machines Corporation.

appearance of superhuman machines in the next few years.

Arguing AI also takes a quick look at other people who have contributed in various ways to the AI debate, including David Hilbert, Alan Turing, Marvin Minsky, Allen Newell, Herbert Simon, Hans Moravec, Daniel Dennett, Hubert Dreyfus, Douglas Hofstadter, and Rod Brooks. The roots of the “AI debate” may draw as heavily on science fiction as on mathematics and philosophy, so the influence of works such as *2001: A Space Odyssey* is considered as well.

This may sound like too much material for such a small book—and it is. *Arguing AI* is extremely superficial. Williams has succeeded pretty well in the task he set for himself: providing intelligent, non-expert readers with a road map for this particular fragment of a much larger debate. But he approaches the task as a journalist: more space is devoted to the personalities than to the ideas, and almost none at all to the technical issues that really lie at the heart of the matter. As much attention is given to Jaron Lanier’s dreadlocks as to the quality of his arguments.

For readers unfamiliar with AI, this book will give them some idea of what all the recent shouting has been about. For those who want to dig deeper, the book has a good bibliography, including both traditional publications and Web sites (though, curiously, the book has no index). For those who want to understand the technical challenges of AI or the intellectual content of the philosophical/social issues surrounding the study of intelligence, this book does not even scratch the surface. But it was fun to read.

Scott E. Fahlman
IBM Thomas J. Watson Research Center
Hawthorne, New York

Note—The books reviewed are those the Editor thinks might be of interest to our readers. The reviews express the opinions of the reviewers.